

Using an Internet-based Decision Research System in Aviation Training Research

Robert Mauro

Department of Psychology
Institute of Cognitive & Decision Sciences
University of Oregon, Eugene OR 97403-1227

Immanuel Barshi

NASA-Ames Research Center
Bldg 262 Room 122
Moffett Field, CA 94035-1000

ABSTRACT

Investigators conducting research on human decision-making in aviation and other applied domains have frequently been deterred from working with people of varying experience, expertise, age, socioeconomic status and culture, because of logistical difficulties. Recent developments in computer technology may provide a solution to many of these problems. Potential participants can be reached at home or at work through the Internet using aesthetically enticing video and audio materials. In this paper, an Internet-based Decision Research System (IDRS) which is capable of reproducing simulated decision environments and conducting research using other typical laboratory procedures is described. This system allows researchers to reach participants at remote locations and gather many of the measures typically obtained in the laboratory (e.g., time spent accessing different sources of information, order of access, self reports of confidence, and subjective probability estimates). The IDRS and other systems like it could dramatically alter the way behavioral research and training program evaluations are conducted.

INTRODUCTION

Human error is implicated in most aviation accidents and incidents. Evidence of errors in judgment and poor decisions can be found littered throughout the reports of most major accidents. Even accidents *not* attributed to human error may be the result of poor decisions. For example, “mechanical failures” are often the result of a poor design or maintenance decision or a poor choice of

materials. Accidents attributed to inadequate pilot skill often result from decisions to enter challenging conditions the pilots knew or should have known would require skills that would exceed their abilities. Thus, understanding decision-making and understanding how to train good decision-making is a prerequisite for substantially improving aviation safety.

Basic research on judgment and decision-making (JDM) has produced an extensive literature.³ Unfortunately, the findings from much of this work may not apply to aeronautical decision-making (ADM).⁴ Although the results of behavioral decision research are often stated in terms that imply broad applicability, this research frequently relies on small samples from special populations. Most of the published research on decision-making is based on the performance of undergraduate student volunteers from universities in the United States working on artificial problems. The flight crews, controllers, and other professionals who make decisions in aviation frequently differ from the typical subjects of JDM research in age, experience, and sometimes culture. ADM problems differ from those typically studied in JDM research on many dimensions, especially complexity and importance. In short, one cannot safely generalize from JDM research to ADM applications.

Given these limitations, it is perhaps not surprising that the JDM literature fails to provide clear direction for programs designed to teach good decision-making skills. There is little systematic research in JDM or ADM on methods for training good decision-making. This is due in part to the complexity of the problem and in

part to the logistical difficulties involved in this research.

To understand how to teach decision-making, in an operational context, one must understand how good decisions are made in that environment. To do so, one must study how individual differences in ability, temperament, and motivation interact with the decision strategy used and the characteristics of the situations encountered to influence the quality of the decisions made. Different situations may call for different decision strategies and individuals may differ in the abilities required to implement these strategies. Hence, it may not be possible to develop a single strategy that can be taught to all individuals to be used in all situations. Instead, a more complex program may be required in which the differences between situations and individuals are recognized.

Multidimensional research of this sort requires large numbers of participants with differing backgrounds. Unfortunately, the logistics of conducting decision research with varied populations can be crippling. For example, in aviation it is often impossible to bring a large and varied sample of pilots to the research laboratory. Even when it is possible, the costs of conducting this research using traditional methods are frequently prohibitive. Pilots' skills can be evaluated from a set of exercises and their declarative knowledge can be measured by multiple choice questions. But to properly evaluate their decision-making abilities, their performance in an appropriate range of situations must be examined. This can be done using Line Oriented Flight Training (LOFT) simulations or check-rides, but these techniques are expensive, time-consuming, and difficult to standardize. Recent developments in computer technology provide another alternative that may solve the logistical problems of this research.

In this paper, we describe an Internet-based Decision Research System (IDRS) that could greatly reduce the difficulties encountered by decision researchers who seek to conduct research with varied populations. This system is capable of reproducing simulated decision environments and conducting research using

other typical laboratory procedures over the Internet. It is designed to gather many of the measures typically obtained in the laboratory (e.g., time spent accessing different sources of information, order in which that information is accessed, self reports of confidence, and subjective probability estimates) at remote locations.

The IDRS is designed to:

- Attract a targeted subject population

- Screen valid subjects

- Provide data security & confidentiality

- Obtain questionnaire data

- Lead subjects through customizable simulated decision-making situations

- Record and store objective and subjective data

- Provide feedback to the subject.

A demonstration project that examines differences in the information-processing and decision-making strategies used by pilots with differing backgrounds and experiences is currently underway.

THE IDRS

Functional Overview

From the perspective of the researcher and participant, the IDRS is composed of three major parts: preliminary questions and materials, decision simulation, and concluding questions and materials.

Preliminary Questions & Materials. This section of the IDRS is devoted to presenting introductory material and obtaining responses to questions. The major components are: an introductory page with links to whatever information the researcher desires (e.g., information about the researchers, the project, and decision research in general), an information page describing the purpose of the research and the procedures to be followed, an informed consent page, and preliminary questionnaire pages that solicit background information. Questions whose answers might be contaminated by previous participation in the simulation could also be asked here. In this section of the demonstration project, the participants are asked about their training, certification, and flight experiences.

Decision Simulation. This section is composed of introductory instructions and a simulated

decision situation through which the participant can navigate. The decision situation is composed of a series of ordered scenes. In each scene the participant has access to potentially different information and must make different decisions. From the participants' perspective, each scene appears to be a domain appropriate environment (e.g., office, cockpit). By using a mouse to point and "click" on an object in the scene (e.g., telephone, manual, communications radio), the participant is able to access the different types of information normally accessed using these objects. This information is presented in pictures, graphs, text, audio, or video. The sources of information accessed, the order of access, and the time spent accessing the information is recorded. The participants' progress is interrupted to obtain information about their cognitive and affective processes (e.g., subjective estimates of the probability of events, self report of emotional state). As in real life, the participants are prevented from returning to earlier scenes within a scenario.



Figure 1. Take-off scene from IDRS Prototype

In the demonstration project, the decision scenario simulates a typical general aviation problem. The participants must decide whether to conduct a flight in a single-engine general aviation aircraft given the weather, their abilities, and the capabilities of the aircraft provided. If they decide to make the flight, the pilots must then decide how to conduct the flight on the ground and at different points along the route (e.g., under visual (VFR) or instrument

(IFR) flight rules and at what altitude). This scenario was based on actual (and fairly typical) winter weather in the Northwest.



Figure 2. Cockpit view from the IDRS prototype.

In the first scene, the participant is placed in the briefing room of a Fixed Based Operator (FBO). All of the information that is typically available to a pilot is provided: a telephone to contact a Flight Service Station (FSS) briefer for a weather report, a computer to access an on-line weather service, the pilot's operating handbook (POH), IFR and VFR charts, and visual cues.

In the following scenes, the participant is placed in the cockpit of the aircraft at different points along the flight (see Figures 1 and 2). Again, most of the information commonly available to the pilot is provided (e.g., flight, navigation, and engine instruments, access to Air Traffic Control (ATC), FSS, and automatic terminal information service (ATIS), charts, POH, and visual cues).

The program records what information is accessed, in what order, and for how long. After the pilots make their decisions, they are asked to provide subjective probability estimates for important possible events and to report on their affective state.

Concluding Questions and Materials. The final portion of the IDRS is devoted to obtaining retrospective reports, evaluations, open-ended comments, and answers to questions not asked in the preliminary section. Once these measures

are obtained, the participant is guided to a final debriefing page with links to other information about the research project and related scientific and professional information.

Upon completion of the scenario in the demonstration project, the participants are asked to answer questions about their decisions and to take a brief personality measure. The pilots are then debriefed and asked to comment on their experience as research participants using the IDRS. Finally, the participants are guided to other Internet sites containing relevant professional and scientific information (e.g., Federal Aviation Administration Human Factors, NASA Aviation Safety, AOPA Aviation Safety Foundation, Society for Judgment & Decision-Making).

Technical Overview

The IDRS program system can be analyzed into two components: (1) A client-side component that provides the simulation experience, gathers the needed data, and transmits it to a server in a secure format, and (2) a server-side component that accepts data from the client, and then validates, processes, and stores the information in a database (see Figure 3). Once a person agrees to participate in the research, a small program is automatically downloaded onto the participant's computer. Once the experiment is complete, the data is uploaded to a web-server and the program erased. The two components of this system are discussed below.

Client-side Component. The client-side component consists of a self-running program that downloads automatically via the Internet. It was developed using Director, Flash, and Java and is compatible with both Macintosh and Windows/Intel computer platforms using most common network software. The client-side is downloaded, while the participant answers the preliminary questions. Download times range from 30 seconds to 5 minutes depending on the complexity of the scenario and the speed of the user's modem. Long scenarios can be downloaded in parts so that the participant is

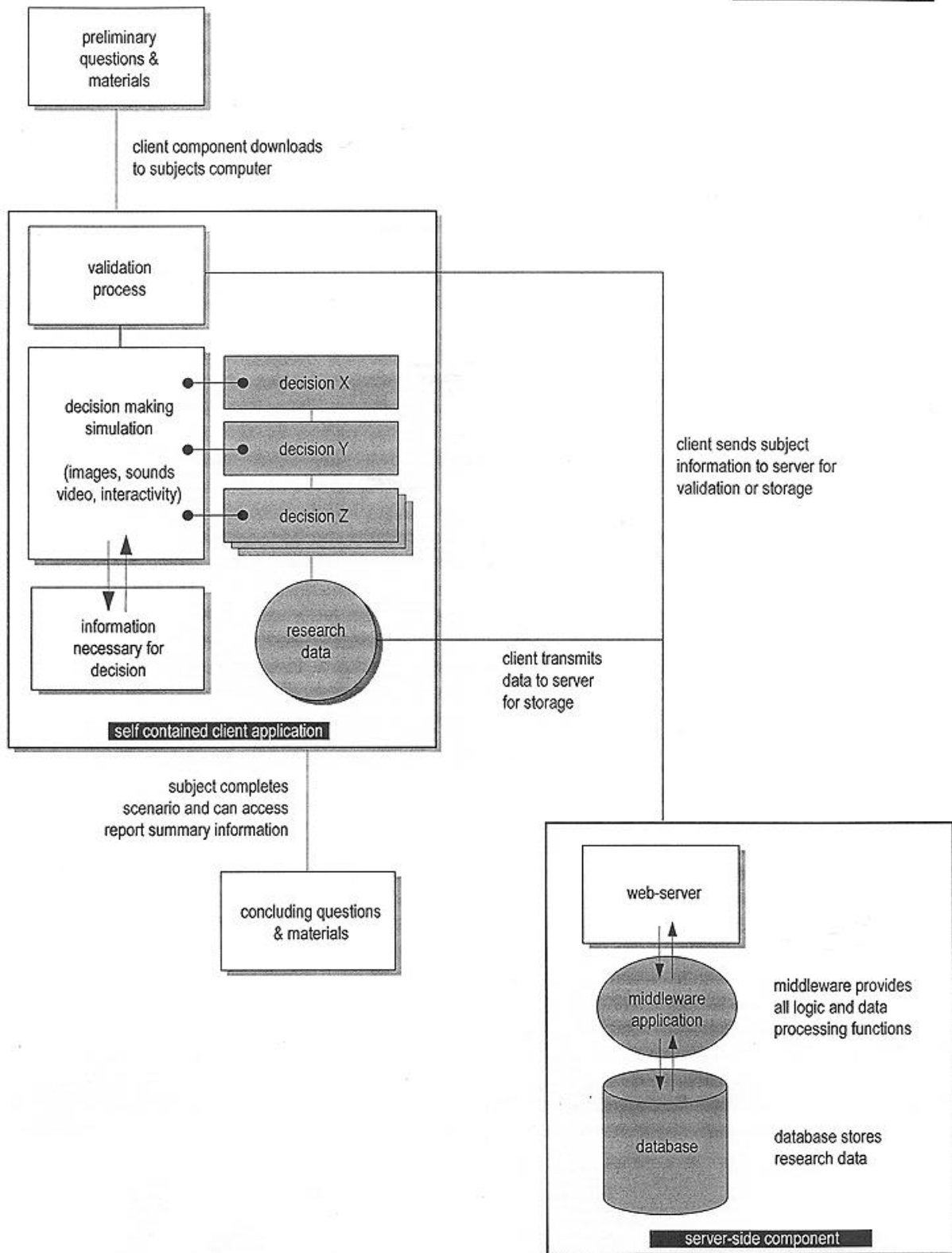
engaged in activities while the client-side continues to download. The client-side runs completely within the web browser (e.g., Netscape, Microsoft Explorer) so no software other than extensions commonly incorporated in web browsers is required. These browser extensions are commonly included with browser packages but can also be downloaded by the IDRS over the Internet at no cost.

Once downloaded, the client-side connects to the server-side and confirms that it is functioning properly. To protect the integrity of the data and the confidentiality of the participants, data are transmitted in an encrypted format using the same security system used in on-line financial transactions. At no time is any data collected by the server publicly accessible. Once the data collection begins, the client-side program has complete control over the data flow. The client-side is capable of time-stamping any action by the participant and maintaining the integrity of the data regardless of network latency.

Although platform differences make reliable millisecond timing impossible, one-tenth of a second accuracy is possible. When the experiment is complete, the data is transferred from the client-side to the server-side via a secure connection and stored on the server. This modular approach affords a higher degree of performance than is possible using traditional web materials. The interactive use of sound, images, and video would not be viable using a conventional content-on-demand system like those commonly used in web development.

Server Component. The server component consists of three independent but tightly integrated packages: a secured web-server, a database application, and a 'middleware' bridge that connects the web-server to the database.

Once a participant agrees to take part in the research, the web-server downloads the client-side software to the participant's computer. The server also allows the client-side software to communicate with the middleware. The middleware provides the needed validation, storage and reporting functions. It acts as a transaction manager between the web-server and the database. Once the middleware verifies that



the data is ready for **Figure 3. Technical**

Overview of IDRS.

storage, it passes the information to the database. The prototype database was developed on a Windows NT server using an SQL compliant software package. This provides adequate scalability and performance to process over 10,000 participants initially with a future capacity well beyond 100,000.

Some Issues with Internet-based Evaluation Research

Even though research conducted over the Internet may solve some of the logistical problems associated with traditional laboratory studies, it is not without its own problems. Using the Internet to conduct research may result in recruiting a biased population. There may be difficulties in recruiting, retaining, and selecting participants. In addition, the researcher may lose some control over the experimental environment.

Contacting Participants

In some cases, obtaining a sample is relatively easy. The potential participants may all be in one place at one time. For example, if the target population is comprised of students enrolled in a college-based training program, arrangements for testing may be made during a class period. In other cases, obtaining an appropriate sample is considerably more difficult. For example, if the target population is comprised of the general aviation (GA) pilots who have purchased a particular videotape course, arranging to test these pilots is a logistical challenge. A similar problem occurs when it is necessary to delay testing of a group that was trained together. For example, pilots for an airline may be trained in a relatively small number of locations but work all across the country or the world. To investigate the long-term effects of the training program or to disguise the connection between the training program and the test situation, researchers may need to wait some time before testing the pilots. Under these circumstances, it may prove

very difficult to arrange the testing sessions. Bringing either the dispersed airline crew members or the scattered GA pilots into a laboratory is impractical.

The IDRS provides a means of bringing much of the laboratory to the subject. If the potential subject has access to a modern personal computer, the researcher using an Internet-based system need only contact the potential subjects and induce them to participate. When the potential subjects are employees, the organization may provide the computer and/or require participation. However, when the potential subjects are not so affiliated, contacting and recruiting the subjects is an additional task.

For individuals to be in the population of subjects that could be reached with an Internet-based research tool, they must have access to a computer that is connected to the Internet and that runs appropriate communication software. The size of this population is quite large. Estimates of the number of people in the United States alone who access the Internet range from 40 to 47 million.¹ Furthermore, the population of users is not limited to those in the United States. Indeed, individuals from almost every country operate World Wide Web (WWW) sites. However, the Internet population is not representative of the general population. Users tend to be male, younger, better educated, and better compensated than the broader population.²

Although there is no definitive accounting of who uses the Internet, the number of aviation professionals (e.g., pilots, controllers, dispatchers, mechanics) that can be reached through the Internet is large and increasing. However, these individuals probably do not comprise a random sample of their professions. They may differ from other members of their profession in psychologically important ways. For example, their familiarity with computers may lead them to process information and to make decisions in systematically different

ways. However, computer literate experts dominate many of these fields. Furthermore, alternative means for reaching these professionals for research, may result in samples that are even more biased. Even if a random sample is contacted, only a small non-random sample may agree to come to the laboratory to participate.

Recruitment

Again, if the potential participants are members of a group over which the researcher has some control, recruitment may not be a problem. In other cases, recruitment is an additional task. Simply having access to the Internet will not guarantee that an individual will participate. When the subjects are not required to participate, the research must be advertised. There are Internet sites operated by many government and professional aviation organizations (e.g., NASA, FAA, AOPA, ALPA) and vendors of professional products. There are on-line bulletin boards and electronic mailing lists (e.g., list-serves). Advertisements distributed through these means are likely to reach a large proportion of the target audience. Alternately, if the names and street or electronic addresses of the members of the target audience are known, they may be contacted directly.

Retention

Reaching the target population is only part of the problem. Once contacted, the potential subjects must be convinced to participate. The gifts and nominal monetary rewards commonly offered to research subjects can be offered to Internet subjects and mailed to them. However, these rewards may not lure most professionals. Professionals decide to participate in most research out of curiosity, interest, or concern for their profession. They will continue to participate if they feel that their contribution is important and appreciated and if the experience is enjoyable. Participation may be encouraged by demonstrating the potential importance of the research to

will soon recognized professional organizations, obtaining endorsements from these organizations, and advertising on their web sites. Participants may be offered summaries of the research findings to be sent electronically on completion of the demonstration project and the opportunity to correspond electronically with the researchers. Participants may also be directed to web sites that link to relevant professional and scientific sources. However, probably the most effective way to encourage participation is to make the experimental task meaningful and enjoyable for the participant. The IDRS site that we developed to study aeronautical decision-making incorporates high quality visual and audio material and provides reasonably realistic decision scenarios. Feedback regarding how other participants or recognized experts responded could also be provided. For example, after making a decision a graph could be presented that portrayed the percentage of previous subjects that chose each option. The decision on whether to include this type of feedback must hinge on each investigator's estimate of its potential for biasing other responses.

Verifying group membership

Researchers are frequently interested in comparing the behaviors of different groups of participants. Typically, group membership is determined by manipulated experimental conditions or questionnaire responses (e.g., level of training, amount of experience). Sometimes group membership is determined by physical location (e.g., presence at a professional convention, military unit). Researchers rarely attempt to verify group membership. In most (but certainly not all) cases, it seems unlikely that the participants would be mistaken or lie about group membership. However, there may be more serious problems in determining group membership on the Internet. The medium provides both widespread access and anonymity. Under

these conditions, the power of social norms that encourage honesty may be diminished. Several observers have noted an apparent is sufficiently intriguing and the simulation reasonably enjoyable, individuals outside the target population may be inclined to participate.

Two tacks can be taken to address these selection issues. Participants could be asked to answer a small number of questions that only members of the target group would be expected to answer correctly and/or participants could be asked to provide information that could be used to verify their membership status. Neither approach is foolproof. Some very knowledgeable out-group members may be able to provide correct responses on a knowledge test and some members of the target group may not. Although level of knowledge may be an important predictor of group membership in many cases, it may not always be a reasonable proxy for group membership. Members of many groups (e.g., pilots) must possess licenses to practice. Hence, their names or license numbers frequently can be used to verify their membership status. However, participants may be hesitant to provide this information. When the potential subjects are participating in conjunction with a particular training program, they may be issued identification numbers. The technical problem of providing reasonable security for the participants' responses is easy to address. Convincing potential subjects that their responses actually are confidential requires more work.

A related problem is the possibility that subjects may want to participate more than once. Investigators using standard research paradigms rarely verify that potential participants have not previously participated in a project. However, the ease of access and anonymity provided by an Internet based research system may encourage participants to revisit a site if the experience was sufficiently enjoyable. To test for this possibility, a record of the electronic address (user name) of each participant could be

decrease in civility when individuals engage in discussions over electronic bulletin boards. If the research maintained (different individuals using the same computer will usually have different user names). To ensure confidentiality, this list may be maintained separately from the participants' data. Researchers could then choose to route the user to a different project, use only the first set of responses from an address, or analyze for changes with experience.

Controlling the Experimental Environment.

Although the IDRS measures the time for which information is displayed, it cannot ensure that the participant is attending to that information while it is being displayed. This problem is not unique to the Internet; it is also encountered in laboratory research. However, it may be exacerbated by the uncontrolled environments in which Internet participants are working. This problem can be alleviated to some extent by asking the participants not to begin the experiment until they anticipate having the time to complete it without interruptions. In addition, by presenting information in relatively small packets, excessive access times can be detected easily. If no input occurs after a reasonable amount of time, a request for a response can be generated and that access time datum flagged as potentially inaccurate.

To study how the experimental environment may affect the participants' responses, two approaches are being pursued. First, pilots in the demonstration project are asked to report on their environment (e.g., home, office) and on the conditions (e.g., quiet, noisy) that they encountered while working on the IDRS. The effects of these factors on the participants' responses can then be analyzed. Second, pilots are being asked to participate in the demonstration project in controlled environments (e.g., research laboratories, flight training schools). The responses of these subjects can be compared to those made by similar subjects

participating in the study in uncontrolled environments.

Conclusions and Future Directions

Preliminary results based on a small sample of pilots who have completed the prototype exercise suggest that the

simulation provides a fairly realistic scenario. All of these pilots rated the scenario as very realistic and stated that they believed that they would have made similar decisions on the ground and in the cockpit. A more thorough evaluation is underway.

There are many possible adaptations of the IDRS. One particularly intriguing possible application of the IDRS is in cross-cultural ADM research. Because neither aviation nor the Internet is limited by national boundaries, the IDRS could prove useful in reducing the difficulty and cost of cross-cultural research. Project development, implementation, and analysis facilities could be centrally located while the participants are anywhere in the world. If desired, multilingual applications could be easily developed.

Another intriguing future direction is in crew research. The IDRS could be easily adapted to simulate the presence of another crew member. More work would be required to actually allow two different individuals to interact from remote locations, but the technology exists for this application as well.

The possible applications of a system like the IDRS are not limited to basic research. The IDRS or a system like it could be used in program or student evaluation. A relatively sophisticated examination of the effectiveness of several different training programs or of the students in a single training program could be conducted from a central location even though the students may be widely dispersed. The examinations could be conducted at any time at the students' convenience. This might prove particularly useful to airlines and corporations with widely distributed personnel.

The IDRS is also a natural complement for Computer Based Instruction (CBI) programs. CBI promises to revolutionize the way training is conducted in aviation. Instead of assembling in traditional

classrooms on a set schedule at a central location, students may learn in small groups or alone at their convenience anywhere that they have access to a computer. Today, the medium of choice for CBI is usually CD-ROM. A single CD-ROM can contain a tremendous amount of information together with an intelligent tutoring program that can tailor the instruction to each student's needs or allow the students to learn according to their own predilections. This type of system is made even more powerful when combined with Internet access. Without leaving the CBI program, the student can access Internet resources to obtain more up-to-date information or to go into more depth on a topic covered in the CD-ROM. In addition, performance evaluations can be conducted over the Internet from a central location.

With systems like the IDRS, these evaluations are no longer limited to simple multiple-choice tests. Reasonable approximations of actual operational competence can be obtained. This type of system may prove particularly valuable for organizations, such as airlines, that must conduct recurrent training with employees who are scattered across the country or across the world. It would also be a useful tool for corporations that do not have internal training programs. Pilots, for example, could engage in recurrent training while waiting for flights from anywhere in the world.

Some problems remain. For example, there is no simple foolproof method for verifying computer-users' identities when they are taking examinations at a remote location. Users can be required to enter security codes or to sign and mail-in verification forms imprinted with codes identifying the examination session. Although similar systems are used by financial and government agencies (e.g., the Internal Revenue Service) to verify identity in other contexts, these procedures can be circumvented. Until less malleable means of verification (e.g., visual or voice identification) become more commonplace,

we may need to rely to some extent on the honor system combined with sporadic in-person examinations.

Research that deepens our understanding of how individuals' experience and personality affect the way they make decisions will aid in the development of improved training materials and safety procedures. By reducing the difficulties encountered in conducting this research, the IDRS and systems like it could dramatically alter the way that behavioral research is conducted.

AUTHOR NOTE

The IDRS was developed in conjunction with the University of Oregon New Media Center under the direction of Mike Holcomb. We also thank Patricia Bruininks, Scott Nemeth, Jim Klein, FSS, Greg Thompson, National Center for Atmospheric Research and all of the pilots and researchers who assisted with this project. The development of the IDRS was supported by grants from the National

Science Foundation and the National Aeronautics and Space Administration. Portions of this paper were previously reported at the Tenth International Symposium on Aviation Psychology at Ohio State University, May 3-6, 1999.

REFERENCES

- Cyberatlas (1997). www.cyberatlas.com/demographics.html.
Internet Demographics Study (1997). ssl.hotsites.net/hotsites/web/background/userdemo.html
3. Mellers, B., Schwartz, A., & Cooke, A. (1998) Judgment and decision making. Annual Review of Psychology, 49, 447-477.
4. Orasanu, J., & Fischer, U. (1997). Finding decisions in natural environments: The view from the cockpit. In C. Zsombok & G. Klein (Eds.). Naturalistic decision making (pp. 343-357). Hillsdale, NJ: Lawrence Erlbaum Associates